

Factorized cross entropy integrated hyperspectral CNN (HSCNet-FACE) for hyperspectral image classification

Pawankumar C. Patil, Shashidhar Sonnad

Department of Electronics and Communication Engineering, Sharnbasva University, Kalaburagi, India

Article Info

Article history:

Received Jun 8, 2024
Revised Nov 24, 2024
Accepted Dec 25, 2024

Keywords:

Botswana
Convolutional neural networks
Factorized cross entropy
Hyperspectral image
Indian Pines

ABSTRACT

The use of hyperspectral image classification algorithms has garnered increasing interest from the scientific community in recent years, especially in the field of geosciences for pattern recognition applications. In order to extract full spectral-spatial characteristics, this study presents feature extraction with hyperspectral CNN (HSCNet), a unique hierarchical neural network architecture. HSCNet can handle computational complexity issues and capture extensive spectral-spatial information with ease. We use factorized cross entropy (FACE) to address the common problem of class imbalance in both experimental and real-world hyperspectral datasets in order to construct an accurate land cover classification system. FACE makes it easier to reconstruct the loss function, which helps to effectively accomplish the goals that have been expressed. We provide a new framework for hyperspectral image (HSI) classification called FACE, which combines components from HSCNet and FACE. Next, we carry out in-depth studies using two different remote sensing datasets: Botswana (BS) and Indian Pines (IP). We compare the effectiveness of different backbone networks in terms of categorization and compare its classification performance under various loss functions. Comparing our suggested classification system against the state-of-the-art end-to-end deep-learning-based techniques, we find encouraging results.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Pawankumar C. Patil
Department of Electronics and Communication Engineering, Sharnbasva University
Kalaburagi, Karnataka, India
Email: pavan.plpas@gmail.com

1. INTRODUCTION

In agriculture, hyperspectral imaging (HSI) has proven to be highly useful for a variety of activities, including disease identification, crop management, yield forecasting, and soil, water, and land condition monitoring [1]-[6]. In agricultural sectors, deep learning approaches have shown great promise as tools that can help farmers make timely and educated decisions. Their ability to automatically extract relevant features gives them an advantage over traditional machine learning (ML) techniques. Deep learning algorithms have made tremendous progress in crop categorization tasks in particular, and numerous efforts are being made to improve these techniques in order to better solve agricultural concerns [7]. Agricultural operations have been completely transformed by the integration of deep learning with HSI, which has improved crop categorization accuracy and efficiency among other related applications. A more profitable and sustainable agriculture industry is promoted by this synergy and the recent strides in artificial intelligence (AI) have facilitated its integration with diverse applications across research and industry landscapes.

Concurrently providing a multitude of spectral and geographical data are hyperspectral images (HSIs). The spatial dimension displays details on the ground surface, whereas the spectral aspect reveals the

distinct spectral curve properties of each pixels, both spectral and spatial information are easily integrated by HSIs [8]. However, they complicate feature extraction and terrain categorization because to high dimensionality and information redundancy [9]. In order to overcome this, HSI spectral pixels are frequently projected onto lower-dimensional feature spaces using dimensionality reduction techniques [10]. Two prominent methods in this field are linear discriminant analysis (LDA) and principal component analysis (PCA), finding the eigenvectors that correspond to the greatest eigenvalues in the original data's covariance matrix is the first step in PCA [11]. In order to achieve feature extraction and dimensionality reduction, the original spectral pixels are projected into an orthogonal subspace formed by these eigenvectors. In contrast, LDA enhances data separability within the subspace by projecting spectral pixels onto a low-dimensional subspace that is tuned for maximizing intra-class scatters and decreasing between-class scatters.

In addition to dimensionality reduction via feature extraction, other data analysis methods can improve spectral differences between various terrains. It has been demonstrated at [12] that derivatives of HSI spectral features capture unique properties of different land-cover classes. In particular, they discovered that first-order spectral differentiation avoids dimensionality problems and increases recognition rates, especially when applied to small sample sizes or poorly quality data. Building on this, spectral first-order differentials were recovered from HSIs in a subsequent work, and dimensionality reduction was achieved through the use of locality Fisher discrimination analysis (LFDA) and locality preserving nonnegative matrix factorization (LPNMF), also the subsequent feature fusion greatly improved classification performance [13].

Fractional differentiation has been popular in spectrum analysis, mostly for predicting element or ion compositions in soil or plants, but it is still not often used in spectral classification [14]. For example, fractional differentiation on soil spectral pixels in visible near-infrared spectroscopy was used in a study Vantaggiato *et al.* [15] to determine the amount of salt and soluble ions in the soil. Fractional differentiation was also used to measure soil organic carbon in another study [16], and in order to get the best SOC estimation accuracy, this required merging spectral parameters obtained from various spectral indices based on fractional differentiation. CNNs have demonstrated incredible skill in the terrain categorization of HSIs in recent years [17]. For example, Bougourzi *et al.* [18] used HRSIs with a one-dimensional convolutional neural network (1D-CNN) that only processed spectral data, ignoring spatial cues. Later, spectral pixels' dimensionality was decreased in a different study [19], after which a two-dimensional convolutional neural network (2D-CNN) was used to extract and classify features, therefore including spatial information from HSIs. There is a significant research gap because fractional differentiation, a useful tool for analyzing data in different frequency ranges, is not widely used in classifying spatial and spectral features. This is despite its success in analyzing soil and plant areas. By using fractional differentiation more often in techniques that classify spatial and spectral data, we could improve accuracy and make these methods more useful in different applications.

In this research, a unique method for classifying HSI is presented, utilizing spatial-spectral properties that are acquired by deep learning approaches. First, a convolutional neural network (CNN) named HSCNet is suggested in order to extract crucial spectral and spatial information for HSI characteristics. In comparison to simple 3D-CNN models, HSCNet reduces computational complexity by concentrating on capturing correlations among nearby bands in the HSI. The results show that when comparing models of similar scale, the HSI classification performance is increased and second, the framework called factorized cross entropy (FACE) is used as a loss function to solve the class-imbalanced character of HSI classification. Numerous studies attest to FACE's efficacy in producing the intended classification results. When compared to current deep-learning-based techniques, the suggested classification framework performs satisfactorily, providing a reliable and effective solution for making land cover categorization more efficient.

2. RELATED WORK

Due to deep learning's remarkable feature extraction skills and several domain triumphs, it has become a potent technique for HSI classification in recent years. Several deep learning techniques have been developed to improve the extraction of spatial and spectral information from HSIs by utilizing convolutional neural networks (CNNs), recurrent neural networks (RNNs), and graph neural networks (GNNs). Certain methods use a two-branch design, extracting spectral and spatial characteristics independently and then merging them using different neural networks [20]. On the other hand, some people choose to collect joint spatial-spectral data using a unified feature extractor, like 3D-CNNs [21]. For example, Hu *et al.* [22] presented a 1D-CNN technique that emphasizes the extraction of spectral features, however this frequently turns out to be insufficient for precise classification. In order to overcome this drawback, Makantasis *et al.* [23] developed a 2D-CNN strategy that focuses exclusively on spatial information learning for the categorization of HSI. But these techniques might not completely take advantage of HSI cubes' intrinsic three-dimensionality. In order to close this disparity, Hamida *et al.* [24] proposed a 3D-CNN technique that improves classification performance by utilizing both spectral and spatial data.

By combining 2D and 3D CNN, a hybrid spectral CNN was created in [25]. In this case, 2D-CNN concentrates on extracting higher-level spatial representations, whereas 3D-CNN helps to capture joint spatial-spectral data early in the process. Other works provide evidence that similar approaches are present in the literature [26]. The channel attention mechanism and its use in HSI classification have been the subject of substantial research in recent years in an effort to improve classification performance even further. A prominent illustration is the squeeze and excitation approach, which simulates inter-channel interdependence to improve overall performance and feature representation capabilities. To improve the representation of spatial-spectral information buried within HSI data, for instance, the incorporation of SE blocks into residual networks was suggested in [27]. Furthermore, some research indicates that handling images in the frequency domain provides direct access to both high and low-frequency components, which may enhance the process of extracting features. Few-shot learning techniques are becoming more popular as a result of the realization that deep learning methods sometimes require a large number of labeled training samples, which can be expensive and time-consuming to annotate manually [28], [29].

As a subclass of deep learning, few-shot learning focuses on teaching models to identify classes of samples with as little as three to five examples per class that have been labeled, this skill is comparable to one that people possess. Transfer learning is the mainstay of modern few-shot learning techniques because there are so many labeled samples accessible and this means using previously learned strategies from large-scale labeled sample data to challenges with sparse labeled data. Meta-learning strategies are used in many few-shot learning approaches to help transmit knowledge effectively and these methods entail creating a large number of tasks that closely resemble the target task and using these created tasks to train a model so that the target task can be easily adapted to. Many techniques designed for the categorization of HSIs have been developed, building on well-established few-shot learning models such as the prototype network [28]. For example, DFSL [30] extracted enhanced spatial-spectral joint features from HSIs by using a 3D residual network as a feature extractor. SSPN [31], on the other hand, used local pattern coding to combine spectral and spatial data while utilizing a 1D-CNN to extract features. Unlike other efforts, HSEMD-Net [32] learned prototype representations for each hyperspectral class by using the earth mover distance rather than the Euclidean distance. Furthermore, for few-shot HSI classification, RL-Net [33] used an interactive RE network [34], which is an enhanced version of the model network.

3. METHOD

We present a novel joint network, called HSCNet, that combines the FACE with a hybrid 3D-2D CNN architecture to achieve HSI categorization. Figure 1 illustrates how our method is designed to utilize spectral-spatial feature maps to extract abstract representations from hierarchical space. Using the hybrid 3D-2D-CNN model, we hope to successfully extract meaningful information from hyperspectral data. Moreover, we utilize the FACE to tackle any difficulties arising from the classification of HSIs, as illustrated in Figure 1. By helping to handle challenging samples, the FACE improves classification performance as a whole.

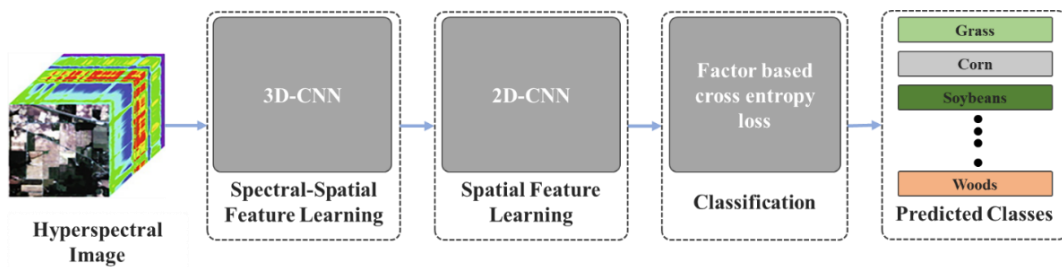


Figure 1. Overview of the proposed feature extractor HSCNet and entire framework HSCNet-FACE

FACE integrated hyperspectral CNN mechanism. The spatial-spectral hyperspectral data cube in this case is represented by the notation $A_{image} \in B^{X \times Y \times S}$, where X , Y , and S stand for the height, width, and spectral bands, respectively, in the HSI. The initial picture, A_{image} has a one-hot label vector $C = \{C_1, C_2, \dots, C_D\} \in B^{1 \times 1 \times D}$, where D is the total number of land cover classes, linked with each pixel. However, the intrinsic complexity of high-dimensional hyperspectral pixels, which frequently display mixes of various land cover classes, presents a considerable barrier in real-world circumstances. Significant spectral intravariability and noticeable similarities between classes are the outcomes of this phenomena.

In order to properly preprocess the data, the image must be segmented into small, overlapping 3D blocks. This segmentation makes it possible to apply our deep learning technique to every single 3D block. The hyperspectral data cube serves as the basis for each neighboring 3D block, which is uniquely recognized by its central spatial coordinates. Every 3D block is represented as $A_e \in B^{f \times f \times S}$, where S is the depth of spectral bands in each block and f is the size of each spatial window. As a result, the 3D blocks are $(X - f + 1) \times (Y - f + 1)$ from the hyperspectral data cube. More specifically, the appropriate spatial window will encompass the height range $[\varepsilon - (f - 1)/2, \varepsilon + (f - 1)/2]$ and the width range $[\delta - (f - 1)/2, \delta + (f - 1)/2]$ for each given hyperspectral 3D block with center spatial coordinates (δ, ε) .

We present the HSCNet-FACE framework, which has a hierarchical convolutional structure intended for categorization of HSI and Gradient descent optimization is used in supervised methods [35] to train the CNNs' parameters. Conventional 2D-CNNs ignore spectral information in favor of constructing discriminative feature maps only over spatial dimensions and as opposed to this, a single 3D-CNN can interact with both spatial and spectral dimensions by traversing the convolutional kernel along all three dimensions (height, width, and spectral). Because of this special feature, the 3D-CNN is able to fully extract spectral and spatial information from the high-dimensional hyperspectral data. HSCNet, the central component of HSCNet-FACE, combines three 3D convolutions, two 2D convolutions, and three fully linked layers in a hierarchical fashion to produce a well-balanced fusion of spectral and spatial information that improves the efficacy of HSI classification.

The output results of 2D-CNN are produced by a convolution method that uses pre-sized 2D filters, also called convolutional kernels, to convolve the input image. The input image's pixels and the filter's weights are multiplied element-by-element in this operation, and the resultant new output pixel value is then computed by summing. Sliding the filter over the whole input image and determining the output pixel value at each place is the convolution process. Through the use of multiple filters at different layers, 2D-CNN is able to efficiently learn a variety of intrinsic properties from the image. The model is then able to perform picture classification and feature extraction tasks by combining these learned features to create higher-level representations. An activation function is used to the features obtained from the convolutional procedure, which adds nonlinearity to the model and the activation value of the r -th feature map in the q -th layer at spatial point (g, h) in 2D convolution is represented by the symbol $m_{(q,r)}^{(g,h)}$, which can be written using (1):

$$m_{(q,r)}^{(g,h)} = \varphi \left(s_{(q,r)} + \sum_{\rho=1}^{n_{l-1}} \sum_{\theta=-\alpha}^{\alpha} \sum_{\mu=-\beta}^{\beta} \kappa_{(q,r,\rho)}^{(\mu,\theta)} \times p_{(q-1,\rho)}^{(g+\mu,h+\theta)} \right) \quad (1)$$

In this case, n_p indicates the number of feature maps in the p th layer, φ stands for the nonlinear activation function, and $s_{q,r}$ for the bias parameter connected to the r -th feature map of the q -th layer. The pre-designed size of the convolutional kernel is represented as $(2\alpha+1) \times (2\beta+1)$ and the weight parameter for the j -th feature map of the i th layer is represented by $\kappa_{q,r}$.

We perform convolutional operations on HSIs using 3D convolutional kernels, as defined in [36] and using discrete or consecutive spectral bands of the input layer, 3D convolutional kernels are applied to create feature maps in the convolutional layer of the HSCNet-FACE architecture. Which efficiently capture spectral information and the activation value of the r -th feature map at spatial location $(g, h, \text{and } t)$ in the q -th layer in 3D convolution, $m_{(q,r)}^{(g,h,t)}$, can be represented as (2):

$$m_{(q,r)}^{(g,h,t)} = \varphi \left(s_{(q,r)} + \sum_{\tau=1}^{d_{l-1}} \sum_{\lambda=-\partial}^{\partial} \sum_{\rho=-\gamma}^{\gamma} \sum_{\sigma=-\delta}^{\delta} \kappa_{(q,r,\rho)}^{(\mu,\theta,\lambda)} \times p_{(q-1,\rho)}^{(g+\mu,h+\theta,t+\lambda)} \right) \quad (2)$$

where the depth of the kernel along the spectral dimension is $2\partial+1$, while other parameters remain consistent with those in (1). We design different 3D convolution kernels in HSCNet with different topologies. $Z \in S_1^w \times S_2^w \times S_3^w$ is the structure of the 3D kernel; w is the layer index of the current kernels, and Z is the number of output channels for the current convolution layer. Z stays constant with the 3D kernel, and the structure of the 2D kernel is similarly indicated as $Z \in S_1^n \times S_2^n$, where n is the layer index of the current kernels. Batch normalization (BN) layers are included as essential elements in our suggested approach and by stabilizing the input data distribution across all network layers, these BN layers are essential in accelerating the training of the overall framework and (3) is the formula for BN:

$$Bn(v) = \frac{v - \bar{v}}{\sqrt{D(v) + \varepsilon}} \cdot i + j \quad (3)$$

The variance in the equation is denoted by $D(v)$ and the mean by \bar{v} . Learnable parameter vectors are indicated by the parameters i and j , whereas ε is a parameter that maintains numerical stability. In order to

provide nonlinear properties to the network, the nonlinear layer is also included and notably, each 3D/2D convolutional layer uses the rectified linear unit [37]. The relative merits of using resampling algorithms versus not using them in HSI classification are up for debate. As a result, we choose to use the Framework FACE as our loss function and it outperforms other traditional loss functions, including entropy loss and multiclass hinge loss, and resolves class imbalance problems more successfully. Adding weighted factors k and $(1 - k)$ for positive and negative classes, respectively, is a commonly used approach to address class imbalance problems. The inverse class frequency, or the reverse of the ratio of positive class samples to negative class samples, is typically used to initialize the parameter α . The goal of this initialization is to balance the class distribution by giving classes with fewer samples during training bigger weights. This is the notation for the k -balanced cross entropy (CE) loss:

$$CE(x, h) = CE(x_t) = -k_t \log(x_t) \quad (4)$$

where x_t is articulated as (5).

$$x_t = \begin{cases} x, & h=1 \\ 1 - x, & h=0 \end{cases} \quad (5)$$

The definition of k in this context is similar to the definition of x_t , where $h \in \{0, 1\}$ denotes the negative classes and positive classes, respectively. It is usually the case that easily classified negatives account for most of the loss [38]. k does not discriminate between easy and hard cases; instead, it strikes a balance between the significance of positive and negative examples. In order to overcome this constraint, the framework FACE reconstructs the balanced cross-entropy loss to down-weight simple examples, shifting the training attention to hard negatives. In (6) is how FACE is formulated:

$$FACE(x_t) = -(1 - x_t)^c \log(x_t) \quad (6)$$

properly, we can use an k -balanced approach of the FACE as (7).

$$FACE(p_t) = -k_t(1 - x_t)^c \log(x_t) \quad (7)$$

4. RESULTS AND DISCUSSION

4.1. Dataset and configuration

The system configuration includes an Intel(R) Core (TM) i7-10750H CPU @ 2.60 GHz, 16 GB RAM, and an NVIDIA GeForce GTX 1650 with 4 GB GPU memory. Python programming and essential packages like Keras, TensorFlow, Spectral, and SciPy are utilized. We perform experiments utilizing four publicly accessible hyperspectral datasets, Indian Pines (IP) and Botswana (BS), to evaluate the efficacy of the suggested technique there are three primary measures are used to assess each method: Kappa coefficient, average accuracy (AA), and overall accuracy (OA). While AA indicates the average accuracy across all categories, OA shows the percentage of correctly identified samples out of all test samples and the degree of agreement between the true values and the categorization outcomes is shown by the Kappa coefficient, here greater classification performance is indicated by higher values of these indicators. 5% of each class from the data was randomly chosen to serve as the training data and the remaining samples as the test data in order to confirm the efficacy of the suggested method's classification.

The IP dataset [39] was collected in 1992 over an area of Indian pine trees in Indiana, USA, using the airborne visible infrared imaging spectrometer (AVIRIS). It consists of 145×145-pixel pictures that span 220 continuous bands at a wavelength between 0.4 and 2.5 μm . Since some bands-the 104th to 108th, 150th to 163rd, and 220th do not reflect water, 200 bands representing 16 different types of ground objects that are used for training. Conversely, the NASA EO-1 satellite collected the BS dataset [39] in May 2001 while it was in the Okavango Delta of BS. This dataset has a spatial resolution of about 20 meters, with dimensions of 1476×256 pixels and a sensor with a wavelength range of 400–2500 nm. Following the removal of noise bands (1 to 9, 56 to 81, 98 to 101, 120 to 133, 165 to 186), 145 bands, which corresponded to 14 different types of ground objects, were used for training.

4.2. Classification results

Among the techniques considered for comparison are the support vector machine (SVM) [40], the spectral-spatial residual network (SSRN) [41], and the fast dense spectral-spatial convolutional network (FDSSC) [42]. A feedback attention-based dense CNN network with a dual-branch multi-attention

mechanism network (DBMA) that makes use of the convolutional block attention module is shown in [43]. Furthermore, a dual-branch dual-attention mechanism network (DBDA) based on the dual attention network (DANet) is introduced in [44]. In order to capture spectral and spatial characteristics, respectively, another technique called dense pyramidal convolution and multi-feature fusion (DPCMF) [45] combines two branches: the spectrum branch and the spatial branch.

Here Figure 2 shows the classification maps for the IP dataset; Figure 2(a) IP single spectral view, Figure 2(b) Ground-truth (GT), and Figure 2(c) GT classes. Furthermore, the Figure 3 shows the OA for different approaches in the IP dataset, where HSCNet-FACE archives 99.71% AA. Here various methodology showcases at x-axis and OA percentage shows at y-axis. Figure 4 shows the Kappa for different approaches in the IP dataset, where HSCNet-FACE archives 99.67% that is 3.35 more compared to top considered model. At Figure 4 various methodology showcases at x-axis and Kappa percentage shows at y-axis.

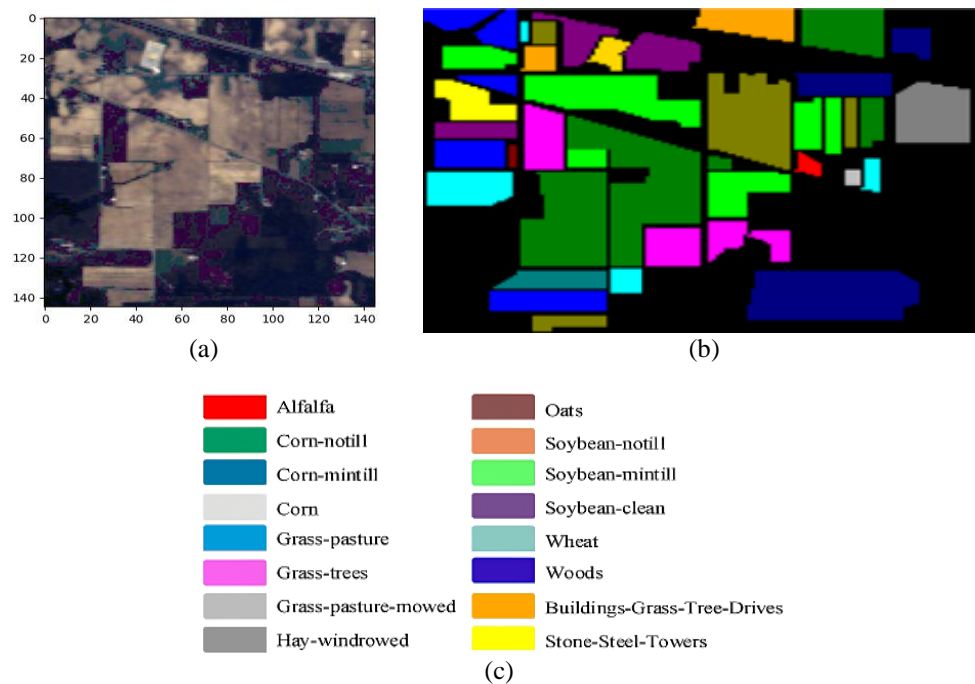


Figure 2. Classification maps for the IP dataset; (a) IP single spectral view, (b) GT, and (c) GT classes

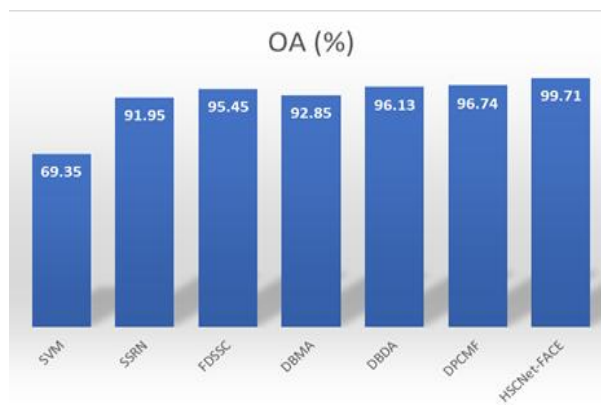


Figure 3. OA for different approaches in the IP dataset

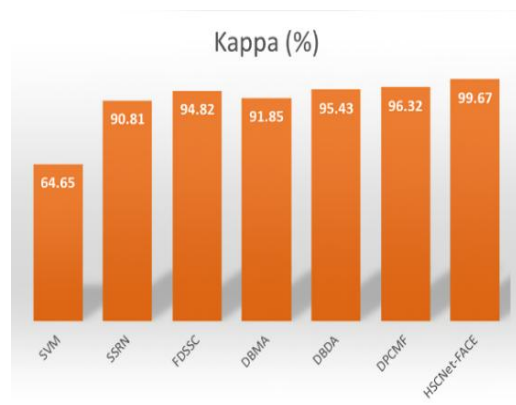


Figure 4. Kappa for different approaches in the IP dataset

Figure 5 displays classification maps for the BS dataset, including Figure 5(a) BS single spectral view, Figure 5(b) GT, and Figure 5(c) GT classes. These maps provide visual representations of the

classification results obtained from different methods. In Figure 6, the OA for various approaches in the BS dataset is presented. Notably, the HSCNet-FACE approach achieves an impressive 99.73% AA, highlighting its superior performance compared to other methods. Here various methodology showcases at x-axis and OA percentage shows at y-axis. Figure 7 illustrates the Kappa coefficients for different approaches applied to the BS dataset. Here various methodology showcases at x-axis and Kappa percentage shows at y-axis. The Kappa coefficient is a measure of agreement between the true values and the classification results. In this context, the HSCNet-FACE approach achieves a Kappa coefficient of 99.71%, which is notably 3.14% higher than the DPCMF approach, indicating its enhanced consistency and accuracy in classifying ground objects. The Tables 1 and 2 displays the classification accuracies (%) of different methods for various classes in a HSI classification task. Each row represents a different class, and each column corresponds to a specific classification method. The last row presents the AA across all classes for each method. Class number indicates the specific class being classified. SVM, SSRN, FDSSC, DBMA, DBDA, and DPCMF: classification accuracies achieved by each respective method for the corresponding class.

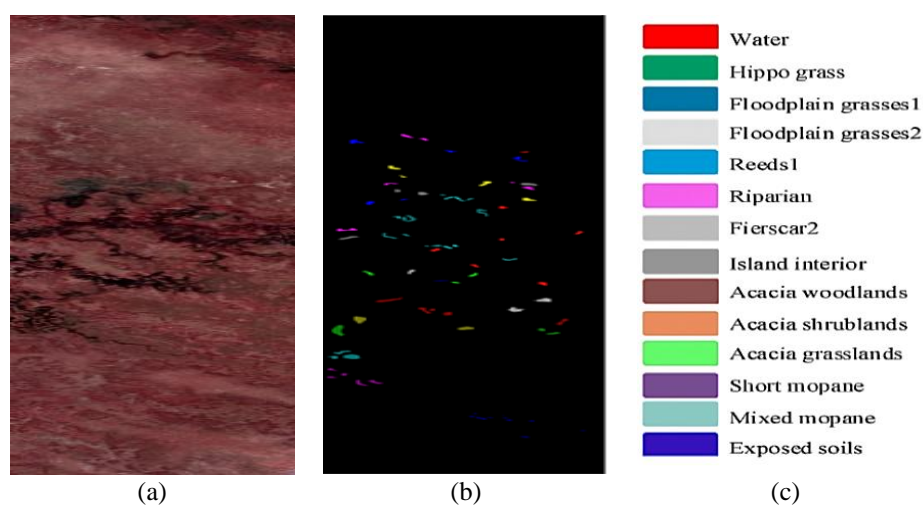


Figure 5. Classification maps for the BS dataset; (a) BS single spectral view, (b) GT, and (c) GT classes

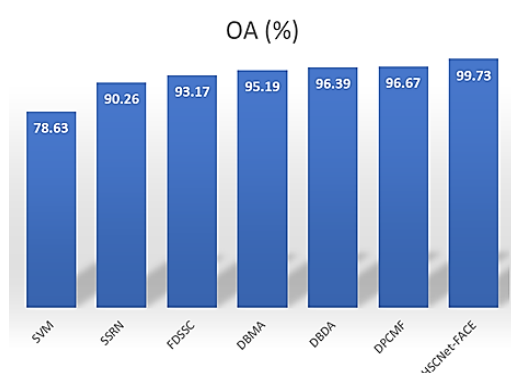


Figure 6. OA for different approaches in the BS dataset

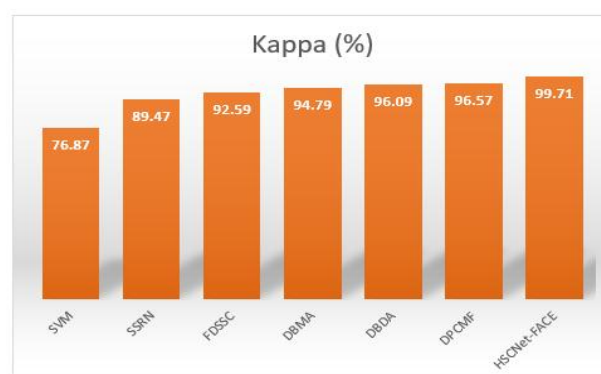


Figure 7. Kappa for different approaches in the BS dataset

The significance of the HSCNet-FACE approach lies in its superior performance compared to other methods across multiple classes. It consistently achieves high classification accuracies, often reaching 100%, indicating its effectiveness in accurately classifying various types of ground objects present in the hyperspectral imagery. The HSCNet-FACE approach achieves the highest overall classification AA (99.73%) compared to other methods at Table 1 and 99.77 at Table 2. It demonstrates superior performance in accurately classifying hyperspectral data across diverse classes. This approach effectively integrates the

HSCNet backbone network with the FACE loss function, enabling the model to capture both spatial and spectral features comprehensively. By leveraging the dual-branch architecture and attention mechanisms, HSCNet-FACE excels in handling the inherent complexities and nuances present in HSIs. Additionally, its robustness is evident in its consistent high accuracy across various classes, showcasing its reliability and effectiveness in practical applications. Overall, the HSCNet-FACE approach stands out as a promising solution for HSI classification tasks, offering state-of-the-art performance and advancing the capabilities in this domain.

Table 1. Categorized results for the IP dataset

Class no.	SVM	SSRN	FDSSC	DBMA	DBDA	DPCMF	HSCNet-FACE
1	24.19	60	97.67	83.33	97.72	100	100
2	56.71	91.47	99.12	92.27	96.43	97.62	99.5
3	65.09	93.51	95.85	92.37	97.81	95.83	99.66
4	39.63	88.95	100	100	97.56	100	100
5	87.33	100	98.35	98.24	98.3	92.86	99.41
6	83.87	95.95	88.44	98.4	96.75	95.24	99.8
7	57.5	86.2	82.75	39.59	88	90.26	100
8	89.28	94.5	100	99.1	100	100	100
9	22.58	69.56	93.33	26.31	100	100	100
10	66.7	84.35	87.31	83.98	91.12	93.1	99.85
11	62.5	91.86	99.09	95.65	98.63	97.26	99.59
12	51.86	86.74	89.01	85.05	93.55	94.12	99.76
13	94.79	98.97	98.92	100	97.97	100	100
14	90.42	94.74	96.24	93.73	94.4	97.3	100
15	62.82	95.09	94.41	94.37	92.98	90.91	99.63
16	98.46	91.11	94.38	96.51	95.45	93.44	98.46
AA (%)	65.86	88.94	94.68	86.18	95.69	96.12	99.73

Table 2. Categorized results for the BS dataset

Class no.	SVM	SSRN	FDSSC	DBMA	DBDA	DPCMF	HSCNet-FACE
1	100	98.47	94.02	97.76	95.97	97.76	100
2	70.7	94.62	100	98.98	98	100	100
3	84.1	87.89	100	100	100	100	100
4	65.95	86.8	96.89	89.4	85.77	88.28	100
5	82.62	74.5	87.5	92.27	93.96	94.36	98.94
6	65.71	80.19	69.76	80.13	87.04	88.65	98.4
7	78.77	90.35	100	96.93	100	99.21	100
8	65.87	87.11	95.6	100	99.32	100	100
9	75.18	93.76	100	94.42	91.04	100	99.55
10	69.82	81.56	91.04	92.77	100	89.7	100
11	95.49	100	100	100	100	100	100
12	93.1	100	88.88	100	100	100	100
13	76.25	96.25	100	100	100	100	100
14	90.41	100	100	97.43	100	100	100
AA (%)	79.57	90.82	94.45	95.72	96.5	97.08	99.77

5. CONCLUSION

The research addresses a gap in HSI classification by proposing the integration of fractional differentiation methods and deep learning techniques. The proposed method, consisting of the HSCNet convolutional neural network and the FACE loss function, effectively captures spatial-spectral features and mitigates class imbalance issues. Results demonstrate improved classification performance compared to traditional methods, offering a reliable and efficient solution for land cover categorization. A comprehensive investigation into HSI classification using various methods and approaches is provided. Through experimentation on datasets such as IP and BS, we have demonstrated the effectiveness of different techniques, including SVM, SSRN, FDSSC, and others. These results highlight the significance of integrating spatial and spectral information effectively, as demonstrated by the HSCNet-FACE approach. Overall, our findings underscore the importance of advanced methodologies in HSI classification and suggest promising avenues for future research in this field. Future research could explore the application of the HSCNet-FACE approach to a wider range of hyperspectral datasets beyond IP and BS. This would help validate its robustness across different environments and land cover types. Developing more efficient versions of the proposed method could enable real-time HSI processing, which is crucial for applications in areas like precision agriculture, environmental monitoring, and disaster management.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Pawankumar C. Patil	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	✓
Shashidhar Sonnad			✓			✓	✓	✓	✓	✓	✓	✓		

C : Conceptualization

M : Methodology

So : Software

Va : Validation

Fo : Formal analysis

I : Investigation

R : Resources

D : Data Curation

O : Writing - Original Draft

E : Writing - Review & Editing

Vi : Visualization

Su : Supervision

P : Project administration

Fu : Funding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

We have obtained informed consent from all individuals included in this study.

DATA AVAILABILITY




Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES




- [1] X. Mei *et al.*, "Spectral-Spatial Attention Networks for Hyperspectral Image Classification," *Remote Sensing*, vol. 11, no. 8, p. 963, Apr. 2019, doi: 10.3390/rs11080963.
- [2] N. Renard, S. Bourennane, and J. B.-Talon, "Denoising and Dimensionality Reduction Using Multilinear Tools for Hyperspectral Images," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 2, pp. 138–142, Apr. 2008, doi: 10.1109/LGRS.2008.915736.
- [3] Y. Dong, B. Du, L. Zhang, and L. Zhang, "Dimensionality Reduction and Classification of Hyperspectral Images Using Ensemble Discriminative Local Metric Learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 5, pp. 2509–2524, May 2017, doi: 10.1109/TGRS.2016.2645703.
- [4] N. Li, D. Zhou, J. Shi, T. Wu, and M. Gong, "Spectral-Location-Spatial Manifold Learning for Hyperspectral Images Dimensionality Reduction," *Remote Sensing*, vol. 13, no. 14, p. 2752, Jul. 2021, doi: 10.3390/rs13142752.
- [5] G. Licciardi, P. R. Marpu, J. Chanussot, and J. A. Benediktsson, "Linear Versus Nonlinear PCA for the Classification of Hyperspectral Data Based on the Extended Morphological Profiles," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 3, pp. 447–451, May 2012, doi: 10.1109/LGRS.2011.2172185.
- [6] T. V. Bandos, L. Bruzzone, and G. C.-Valls, "Classification of Hyperspectral Images With Regularized Linear Discriminant Analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 3, pp. 862–873, Mar. 2009, doi: 10.1109/TGRS.2008.2005729.
- [7] J. Bao, M. Chi, and J. A. Benediktsson, "Spectral Derivative Features for Classification of Hyperspectral Remote Sensing Images: Experimental Evaluation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 2, pp. 594–601, Jun. 2013, doi: 10.1109/JSTARS.2013.2237758.
- [8] P. K. Sathy, C. Pandey, Y. K. Sahu, and S. K. Behera, "Hyperspectral imagery applications for precision agriculture - a systemic survey," *Multimedia Tools and Applications*, vol. 81, no. 2, pp. 3005–3038, Jan. 2022, doi: 10.1007/s11042-021-11729-8.
- [9] B. Li *et al.*, "Above-ground biomass estimation and yield prediction in potato by using UAV-based RGB and hyperspectral imaging," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 162, pp. 161–172, Apr. 2020, doi: 10.1016/j.isprsjprs.2020.02.013.
- [10] A.-K. Mahlein *et al.*, "Development of spectral indices for detecting and identifying plant diseases," *Remote Sensing of Environment*, vol. 128, pp. 21–30, Jan. 2013, doi: 10.1016/j.rse.2012.09.019.
- [11] Ö. Gürsoy, A. C. Birdal, F. Özyonar, and E. Kasaka, "Determining and Monitoring the Water Quality of Kizilirmak River of Turkey: First Results," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XL-7/W3, pp. 1469–1474, Apr. 2015, doi: 10.5194/isprsarchives-XL-7-W3-1469-2015.
- [12] A. Bouguettaya, H. Zarzour, A. Kechida, and A. M. Taberkit, "Deep learning techniques to classify agricultural crops through UAV imagery: a review," *Neural Computing and Applications*, vol. 34, no. 12, pp. 9511–9536, Jun. 2022, doi: 10.1007/s00521-022-07104-9.
- [13] J. F. I. Ntumbirwe and U. L. Opara, "Machine learning applications to non-destructive defect detection in horticultural

- products,” *Biosystems Engineering*, vol. 189, pp. 60–83, Jan. 2020, doi: 10.1016/j.biosystemseng.2019.11.011.
- [14] A. Signoroni, M. Savardi, A. Baronio, and S. Benini, “Deep Learning Meets Hyperspectral Image Analysis: A Multidisciplinary Review,” *Journal of Imaging*, vol. 5, no. 5, p. 52, May 2019, doi: 10.3390/jimaging5050052.
 - [15] E. Vantaggiato, E. Paladini, F. Bougourzi, C. Distante, A. Hadid, and A. T.-Ahmed, “COVID-19 Recognition Using Ensemble-CNNs in Two New Chest X-ray Databases,” *Sensors*, vol. 21, no. 5, p. 1742, Mar. 2021, doi: 10.3390/s21051742.
 - [16] G. Jaiswal, A. Sharma, and S. K. Yadav, “Critical insights into modern hyperspectral image applications through deep learning,” *WIREs Data Mining and Knowledge Discovery*, vol. 11, no. 6, Nov. 2021, doi: 10.1002/widm.1426.
 - [17] F. Bougourzi, C. Distante, F. Dornaika, and A. T.-Ahmed, “CNR-IEMN-CD and CNR-IEMN-CSD Approaches for Covid-19 Detection and Covid-19 Severity Detection from 3D CT-scans,” *Computer Vision – ECCV 2022 Workshops (ECCV 2022)*, 2023, pp. 593–604, doi: 10.1007/978-3-031-25082-8_40.
 - [18] F. Bougourzi, F. Dornaika, and A. T.-Ahmed, “Deep learning based face beauty prediction via dynamic robust losses and ensemble regression,” *Knowledge-Based Systems*, vol. 242, p. 108246, Apr. 2022, doi: 10.1016/j.knsys.2022.108246.
 - [19] B. Lu, P. Dao, J. Liu, Y. He, and J. Shang, “Recent Advances of Hyperspectral Imaging Technology and Applications in Agriculture,” *Remote Sensing*, vol. 12, no. 16, p. 2659, Aug. 2020, doi: 10.3390/rs12162659.
 - [20] F. Zhou, R. Hang, Q. Liu, and X. Yuan, “Hyperspectral image classification using spectral-spatial LSTMs,” *Neurocomputing*, vol. 328, pp. 39–47, Feb. 2019, doi: 10.1016/j.neucom.2018.02.105.
 - [21] B. Fang, Y. Li, H. Zhang, and J. C.-W. Chan, “Collaborative learning of lightweight convolutional neural network and deep clustering for hyperspectral image semi-supervised classification with limited training samples,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 161, pp. 164–178, Mar. 2020, doi: 10.1016/j.isprsjprs.2020.01.015.
 - [22] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li, “Deep Convolutional Neural Networks for Hyperspectral Image Classification,” *Journal of Sensors*, vol. 2015, pp. 1–12, 2015, doi: 10.1155/2015/258619.
 - [23] K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis, “Deep supervised learning for hyperspectral data classification through convolutional neural networks,” in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Jul. 2015, pp. 4959–4962, doi: 10.1109/IGARSS.2015.7326945.
 - [24] A. B. Hamida, A. Benoit, P. Lambert, and C. B. Amar, “3-D Deep Learning Approach for Remote Sensing Image Classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 8, pp. 4420–4434, Aug. 2018, doi: 10.1109/TGRS.2018.2818945.
 - [25] S. K. Roy, G. Krishna, S. R. Dubey, and B. B. Chaudhuri, “HybridSN: Exploring 3-D–2-D CNN Feature Hierarchy for Hyperspectral Image Classification,” *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 2, pp. 277–281, Feb. 2020, doi: 10.1109/LGRS.2019.2918719.
 - [26] M. Q. Alkhatib *et al.*, “Tri-CNN: A Three Branch Model for Hyperspectral Image Classification,” *Remote Sensing*, vol. 15, no. 2, p. 316, Jan. 2023, doi: 10.3390/rs15020316.
 - [27] L. Wang, J. Peng, and W. Sun, “Spatial–Spectral Squeeze-and-Excitation Residual Network for Hyperspectral Image Classification,” *Remote Sensing*, vol. 11, no. 7, p. 884, Apr. 2019, doi: 10.3390/rs11070884.
 - [28] J. Snell, K. Swersky, and R. Zemel, *Prototypical Networks for Few-shot Learning*, vol. 30, 2017.
 - [29] A. Li, W. Huang, X. Lan, J. Feng, Z. Li, and L. Wang, “Boosting Few-Shot Learning With Adaptive Margin Loss,” in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2020, pp. 12573–12581, doi: 10.1109/CVPR42600.2020.01259.
 - [30] B. Liu, X. Yu, A. Yu, P. Zhang, G. Wan, and R. Wang, “Deep Few-Shot Learning for Hyperspectral Image Classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 4, pp. 2290–2304, Apr. 2019, doi: 10.1109/TGRS.2018.2872830.
 - [31] H. Tang, Y. Li, X. Han, Q. Huang, and W. Xie, “A Spatial–Spectral Prototypical Network for Hyperspectral Remote Sensing Image,” *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 1, pp. 167–171, Jan. 2020, doi: 10.1109/LGRS.2019.2916083.
 - [32] J. Sun, X. Shen, and Q. Sun, “Hyperspectral Image Few-Shot Classification Network Based on the Earth Mover’s Distance,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2022, doi: 10.1109/TGRS.2022.3191541.
 - [33] X. Ma, S. Ji, J. Wang, J. Geng, and H. Wang, “Hyperspectral Image Classification Based on Two-Phase Relation Learning Network,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 12, pp. 10398–10409, Dec. 2019, doi: 10.1109/TGRS.2019.2934218.
 - [34] F. Sung, Y. Yang, L. Zhang, T. Xiang, P. H. S. Torr, and T. M. Hospedales, “Learning to Compare: Relation Network for Few-Shot Learning,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun. 2018, pp. 1199–1208, doi: 10.1109/CVPR.2018.00131.
 - [35] S. Srinivasan *et al.*, “Deep Convolutional Neural Network Based Image Spam Classification,” in *2020 6th Conference on Data Science and Machine Learning Applications (CDMA)*, Mar. 2020, pp. 112–117, doi: 10.1109/CDMA47397.2020.00025.
 - [36] S. N. Boualia and N. E. B. Amara, “3D CNN for Human Action Recognition,” in *2021 18th International Multi-Conference on Systems, Signals & Devices (SSD)*, Mar. 2021, pp. 276–282, doi: 10.1109/SSD52085.2021.9429429.
 - [37] Y. Li, W. Wei, J. Zhang, L. Zhang, and Y. Zhang, “Data-Specific Activation Function Learning for Hyperspectral Image Classification,” in *2020 8th International Conference on Orange Technology (ICOT)*, Dec. 2020, pp. 1–4, doi: 10.1109/ICOT51877.2020.9468782.
 - [38] L. Li, M. Doroslovacki, and M. H. Loew, “Approximating the Gradient of Cross-Entropy Loss Function,” *IEEE Access*, vol. 8, pp. 111626–111635, 2020, doi: 10.1109/ACCESS.2020.3001531.
 - [39] M. Graña, M. Veganzons, and B. Ayerdi, “Hyperspectral Remote Sensing Scenes,” *Grupo de Inteligencia Computacional de la Universidad del País Vasco*, 2018.
 - [40] M. Pal, “Ensemble of support vector machines for land cover classification,” *International Journal of Remote Sensing*, vol. 29, no. 10, pp. 3043–3049, May 2008, doi: 10.1080/01431160802007624.
 - [41] Z. Zhong, J. Li, Z. Luo, and M. Chapman, “Spectral–Spatial Residual Network for Hyperspectral Image Classification: A 3-D Deep Learning Framework,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 2, pp. 847–858, Feb. 2018, doi: 10.1109/TGRS.2017.2755542.
 - [42] W. Wang, S. Dou, Z. Jiang, and L. Sun, “A Fast Dense Spectral–Spatial Convolution Network Framework for Hyperspectral Images Classification,” *Remote Sensing*, vol. 10, no. 7, p. 1068, Jul. 2018, doi: 10.3390/rs10071068.
 - [43] W. Ma, Q. Yang, Y. Wu, W. Zhao, and X. Zhang, “Double-Branch Multi-Attention Mechanism Network for Hyperspectral Image Classification,” *Remote Sensing*, vol. 11, no. 11, p. 1307, Jun. 2019, doi: 10.3390/rs11111307.
 - [44] R. Li, S. Zheng, C. Duan, Y. Yang, and X. Wang, “Classification of Hyperspectral Image Based on Double-Branch Dual-Attention Mechanism Network,” *Remote Sensing*, vol. 12, no. 3, p. 582, Feb. 2020, doi: 10.3390/rs12030582.
 - [45] J. Zhang *et al.*, “Hyperspectral Image Classification Based on Dense Pyramidal Convolution and Multi-Feature Fusion,” *Remote Sensing*, vol. 15, no. 12, p. 2990, Jun. 2023, doi: 10.3390/rs15122990.

BIOGRAPHIES OF AUTHORS

Pawankumar C. Patil    is working as an Assistant Professor in ECE Department in Faculty of Engineering and Technology (Co-education), Sharnbasva University, Kalaburagi, Karnataka, India. He has completed B.E. (ECE), M.Tech. (Digital Electronics), and pursuing Ph.D. (Image Processing), IEEE student member. Having more than 10 years of teaching experience and 3 years in research. He can be contacted at email: pavan.plpas@gmail.com.



Shashidhar Sonnad    is currently working as a Professor and Chairman in the Department of Electronics and Communication Engineering, Faculty of Engineering and Technology (Co-Education), Sharnbasva University, Kalaburagi, Karnataka, India. He has 17 years of experience in teaching and 3 years of experience in research. He has completed his M.Tech. (Communication Systems) and Ph.D. (Geospatial Image Processing). He contributed 08-research papers in the International Journals, attended 8-National and International Conferences and 7-papers presented in the conferences. He worked as a reviewer for various International Journals and IEEE sponsored International Conferences. He can be contacted at email: Shashidharsonnad1@gmail.com.